PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/790,043A*

Sector1819

DATE: 10/22/97
TIME: 12:56:35

INPUT SET: S21089.raw

#7/ウR
07/04/98

This Raw Listing contains the General Information Section and up to the first 5 pages.

~~ENTERED~~

```
 1                              SEQUENCE LISTING
 2
 3     (1)      General Information
 4
 5              (i) APPLICANT: Lonsdale, John
 6                             Milner, Peter
 7                             Payne, David
 8                             Pearson, Stewart
 9
10              (ii) TITLE OF THE INVENTION: Novel FabI
11
12              (iii) NUMBER OF SEQUENCES: 2
13
14              (iv) CORRESPONDENCE ADDRESS:
15                 (A) ADDRESSEE: SmithKline Beecham Corporation
16                 (B) STREET: 709 Swedeland Road
17                 (C) CITY: King of Prussia
18                 (D) STATE: PA
19                 (E) COUNTRY: USA
20                 (F) ZIP: 19406-0939
21
22              (v) COMPUTER READABLE FORM:
23                 (A) MEDIUM TYPE: Diskette
24                 (B) COMPUTER: IBM Compatible
25                 (C) OPERATING SYSTEM: DOS
26                 (D) SOFTWARE: FastSEQ for Windows Version 2.0
27
28              (vi) CURRENT APPLICATION DATA:
29                 (A) APPLICATION NUMBER:
30                 (B) FILING DATE: 28-August-1997
31                 (C) CLASSIFICATION:
32
33              (vii) PRIOR APPLICATION DATA:
34                 (A) APPLICATION NUMBER: 60/024845
35                 (B) FILING DATE: 28-AUG-1996
36
37              (viii) ATTORNEY/AGENT INFORMATION:
38                 (A) NAME: Gimmi, Edward R
39                 (B) REGISTRATION NUMBER: 38,891
40                 (C) REFERENCE/DOCKET NUMBER: GM50005
41
42              (ix) TELECOMMUNICATION INFORMATION:
43                 (A) TELEPHONE: 610-270-4478
44                 (B) TELEFAX: 610-270-5090
45                 (C) TELEX:
46
```

```
47              (2) INFORMATION FOR SEQ ID NO:1:
48
49          (i) SEQUENCE CHARACTERISTICS:
50             (A) LENGTH: 256 amino acids
51             (B) TYPE: amino acid
52             (C) STRANDEDNESS: single
53             (D) TOPOLOGY: linear
54
55          (ii) MOLECULE TYPE: protein
56
57          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
58
59   Met Leu Asn Leu Glu Asn Lys Thr Tyr Val Ile Met Gly Ile Ala Asn
60    1               5                  10                  15
61   Lys Arg Ser Ile Ala Phe Gly Val Ala Lys Val Leu Asp Gln Leu Gly
62               20                  25                  30
63   Ala Lys Leu Val Phe Thr Tyr Arg Lys Glu Arg Ser Arg Lys Glu Leu
64           35                  40                  45
65   Glu Lys Leu Leu Glu Gln Leu Asn Gln Pro Glu Ala His Leu Tyr Gln
66       50                  55                  60
67   Ile Asp Val Gln Ser Asp Glu Glu Val Ile Asn Gly Phe Glu Gln Ile
68   65                  70                  75                  80
69   Gly Lys Asp Val Gly Asn Ile Asp Gly Val Tyr His Ser Ile Ala Phe
70                   85                  90                  95
71   Ala Asn Met Glu Asp Leu Arg Gly Arg Phe Ser Glu Thr Ser Arg Glu
72               100                 105                 110
73   Gly Phe Leu Leu Ala Gln Asp Ile Ser Ser Tyr Ser Leu Thr Ile Val
74           115                 120                 125
75   Ala His Glu Ala Lys Lys Leu Met Pro Glu Gly Gly Ser Ile Val Ala
76       130                 135                 140
77   Thr Thr Tyr Leu Gly Gly Glu Phe Ala Val Gln Asn Tyr Asn Val Met
78   145                 150                 155                 160
79   Gly Val Ala Lys Ala Ser Leu Glu Ala Asn Val Lys Tyr Leu Ala Leu
80               165                 170                 175
81   Asp Leu Gly Pro Asp Asn Ile Arg Val Asn Ala Ile Ser Ala Gly Pro
82               180                 185                 190
83   Ile Arg Thr Leu Ser Ala Lys Gly Val Gly Gly Phe Asn Thr Ile Leu
84           195                 200                 205
85   Lys Glu Ile Glu Glu Arg Ala Pro Leu Lys Arg Asn Val Asp Gln Val
86       210                 215                 220
87   Glu Val Gly Lys Thr Ala Ala Tyr Leu Leu Ser Asp Leu Ser Ser Gly
88   225                 230                 235                 240
89   Val Thr Gly Glu Asn Ile His Val Asp Ser Gly Phe His Ala Ile Lys
90               245                 250                 255
91
92              (2) INFORMATION FOR SEQ ID NO:2:
93
94          (i) SEQUENCE CHARACTERISTICS:
95             (A) LENGTH: 771 base pairs
96             (B) TYPE: nucleic acid
97             (C) STRANDEDNESS: double
98             (D) TOPOLOGY: linear
99
```

INPUT SET: S21089.raw

```
100          (ii)  MOLECULE TYPE: Genomic DNA
101
102          (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:2:
103
104    ATGTTAAATC TTGAAAACAA AACATATGTC ATCATGGGAA TCGCTAATAA GCGTAGTATT      60
105    GCTTTTGGTG TCGCTAAAGT TTTAGATCAA TTAGGTGCTA AATTAGTATT TACTTACCGT     120
106    AAAGAACGTA GCCGTAAAGA GCTTGAAAAA TTATTAGAAC AATTAAATCA ACCAGAAGCG     180
107    CACTTATATC AAAATTGATGT TCAAAGCGAT GAAGAGGTTA TTAATGGTTT TGAGCAAATT     240
108    GGTAAAGATG TTGGCAATAT TGATGGTGTA TATCATTCAA TCGCATTTGC TAATATGGAA     300
109    GACTTACGCG GACGCTTTTC TGAAACTTCA CGTGAAGGCT TCTTGTTAGC TCAAGACATT     360
110    AGTTCTTACT CATTAACAAT TGTGGCTCAT GAAGCTAAAA AATTAATGCC AGAAGGTGGT     420
111    AGCATTGTTG CAACAACATA TTTAGGTGGC GAATTCGCAG TTCAAAATTA TAATGTGATG     480
112    GGTGTTGCTA AAGCGAGCTT AGAAGCAAAT GTTAAATATT TAGCATTAGA CTTAGGTCCT     540
113    GATAATATTC GCGTTAATGC AATTTCAGCT GGTCCAATCC GTACATTAAG TGCAAAAGGT     600
114    GTGGGTGGTT TCAATACAAT TCTTAAAGAA ATCGAAGAGC GTGCACCTTT AAAACGTAAC     660
115    GTTGATCAAG TAGAAGTAGG TAAAACAGCG GCTTACTTRT TAAGTGACTT ATCAAGTGGC     720
116    GTTACAGGTG AAAATATTCA TGTAGATAGC GGATTCCACG CAATTAAATA A             771
117
118
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION   *US/08/790,043A*

INPUT SET: S21089.raw

Line	Error	Original Text